# Cloning and Sequence Analysis of a Wheat RING Domain Ubiquitin Protein Ligase Gene

Zaijun Yang & Zhengsong Peng (Corresponding author)
Key Laboratory of Crop Genetic Resources and Improvement Ministry of Education
Sichuan Agricultural University, Yaan, Sichuan, 625014, China
College of Life Science, China West Normal University
Nanchong, Sichuan, 637009, China
Tel: 86-817-256-8011 E-mail: pzs8833@163.com

Kai Wu & Shuhong Wei
College of Life Science, China West Normal University
Nanchong, Sichuan, 637009, China

Yonghong Zhou (Corresponding author)
Key Laboratory of Crop Genetic Resources and Improvement Ministry of Education
Sichuan Agricultural University, Yaan, Sichuan, 625014, China
Tel: 86-28-826-91005 E-mail: zhouyh@sicau.edu.cn

Received: March 21, 2011 Accepted: April 8, 2011 doi:10.5539/mas.v5n3p164

*This work was supported by National Natural Science Foundation of China. No. 30871533*

**Abstract**

The open reading frame (ORF) of RING domain ubiquitin protein ligase gene TaZFP-1 was acquired by RT-PCR methods. Then, analysis and prediction on the acquired sequences and its amino acids, physicochemical properties, hydrophobicity/hydrophilicity, secondary structure, functional domains, sequence alignment and phylogenetic tree. The results showed that TaZFP-1 gene cDNA was 759 bp in length, encoding 252 amino acids (GenBank accession No. JF727656). Most amino acids in TaZFP-1 protein are hydrophilic amino acids, so the protein may be a soluble protein. Secondary structure of TaZFP-1 was mainly composed of α-helices and random coils. Functional domains analysis indicated that the TaZFP-1 is a RING finger domain protein and containing a C3HC4 motif. The molecular evolution trees showed that the TaZFP-1 was clustered into the monocotyledon group and high genetic relationship with *O. sativa* RING domain E3 ligases.

**Keywords:** RING domain ubiquitin protein ligase, TaZFP-1, Gene clone, Wheat

## 1. Introduction

The ubiquitin-proteasome pathway involves multiple steps and requires at least three proteins: the ubiquitin activating enzyme (E1), ubiquitin-conjugating enzyme (E2) and ubiquitin protein ligase (E3) (Weissman, 2001; VanDemark et al, 2002; Passmore et al, 2004). The importance of the ubiquitin-proteasome pathway to cellular regulation in eukaryotes has become increasingly apparent during the last several years. This fact was formally acknowledged recently by the awarding of the 2004 Nobel Prize in Chemistry to Aaron Ciechanover, Avram Hershko, and Irwin Rose for the discovery of ubiquitin-mediated protein degradation. In plants, regulated protein degradation by the ubiquitin-proteasome contributes significantly to development by affecting a wide range of processes, including embryogenesis, hormone signaling, and senescence (Moon et al, 2004).

The E3 ubiquitin ligases comprise a large and diverse family of proteins or protein complexes containing either a HECT domain or a RING domain. HECT E3s are large proteins, typically 100~400KD and existing widely in higher plant. But up to now, the function and substrate of the HECT E3s are unclearly. The single subunit RING

E3s is characterized by the presence of a zinc binding motif, or RING finger. In molecular biology, a RING (Really Interesting New Gene) finger domain is a protein structural domain of zinc finger type which contains a Cys3-His-Cys4 (C3HC4) amino acid motif which binds two zinc cations (Borden et al, 1996; Hanson et al, 1991; Freemont et al, 1991; Lovering et al, 1993).

The common wheat (*T. aestivum* L.) line Three Pistils (TP), which selected by Peng (2003) from the 'trigrain' wheat variety, is a valuable mutant for the wheat breeding. Because the TP mutation has normal spike morphology but produces 3 pistils per floret, it has potential to increase grain number per spike.

In this study, an EST sequence was isolated from SSH cDNA library of wheat three pistils lines (Yang et al, 2011). Blastx searches of the sequences against the GeneBank database revealed that the EST is homology to maize RING ubiquitin protein ligase, we tentatively designated the EST as TaZFP-1. Using RT-PCR method, we obtain the open reading frame (ORF) of the TaZFP-1. We then analyzed the sequence characteristics of the protein encoded by the cDNA and compared it with other plants RING ubiquitin protein ligase reported.

## 2. Materials and Methods

### *2.1 Plant materials*

The common wheat line three pistils (TP) was developed in our laboratory (Peng, 2003), the young spikes at reproductive stage (spike size about 10mm) were collected for RNA extraction.

### *2.2 RNA isolation*

Fresh young spikes were used for total RNA isolation according to the method of Manickavelu (2007). Briefly, in 1.5 ml tube by giving liquid nitrogen as cooling jacket, sample was grinded with glass rod. Modified extraction buffer was used and lithium chloride treatment was done twice to get purified RNA. DNAse treatment was done before proceeding to cDNA synthesis. Quality of RNA was confirmed by gel electrophoresis and concentration was calculated after measured in 260/280nm of spectrophotometer.

### *2.3 Primer design, RT-PCR, cloning of RT-PCR products, and sequencing*

The 553bp EST sequence which isolated from SSH cDNA library of wheat three pistils lines acting as the information probe was sent into the EST database of *T. aestivum* L. to do the BLAST searching (Expect = 10), and then the generated contigs with high identity to the probe and coming from the same organism, *T. aestivum* L., was collected and assembled into a longer novel EST sequences as second probe. The above step was not repeated until the newly generated probe can not be elongated. The polymerase chain reaction (PCR) primers were designed by Primer Primer 5.0, based on the sequence of newly generated contig. The specific primers were as follows: TaZFP-1 forward: 5`-ACAAAGCTTGACCTTAACCATCC-3`; TaZFP-1 reverse: 5`-CGAATTCATGAGTAATACCAGCG-3`.

Total RNAs were synthesized into the first-stranded cDNAs using a reverse transcription kit with Oligo dT as the primers according to manufacturer instructions (Kit cat number: M5101, Promega, USA). Twenty microliters of the first-strand cDNA synthesis reaction system was included in 1 mg total RNAs, 5 mM $MgCl_2$, 1 mM dNTPs, 0.5 mg Oligo $dT_{15}$, 10 U/mL RNase inhibitor, and 15 U AMV reverse transcriptase, and incubated at 42°C for 60 min.

The first-strand cDNA synthesized was used as a template. The total reaction volume for DNA amplification was 25 μL. Reaction mixtures contained 1.5 mM $MgCl_2$, 200 μM of each of dATP, dGTP, dCTP and dTTP (Omega, China), 0.3 μM of each primer, 5.0 units Taq plus DNA polymerase (TaKaRa, Dalian, China). DNA amplification was performed using an MJ Research thermocycler, Model PTC-200 (Watertown, MA, USA) with a program of 5 min at 94°C, followed by 35 cycles of 30 sec at 94°C , 30 sec at 60°C and 1min at 72°C, and then ended with the final extension for 10 min at 72°C. After amplification, PCR products were separated by electrophoresis on 1.5% agarose gel with 1× TAE (Tris-acetate-EDTA) buffer, stained with ethidium bromide and visualized under UV light. The expected fragments of PCR products were harvested and purified from gel using a DNA harvesting kit (Omega), and then ligated into a pMD19-T vector at 4°C for 12 h. The recombinant molecules were transformed into *E. coli* competent cells (DH5α), and then spread on an LB-plate containing 50μg/mL ampicillin, 200mg/mL IPTG (isopropyl-beta-D-thiogalactopyranoside), and 20mg/ mL X-gal. Plasmid DNA was isolated and digested by *Pst*I and *Sca*II to verify the insert size. Five positive clones were selected randomly and sequenced by Huada Zhongsheng Scientific Corporation (Shenzheng, China).

### *2.4 Data analysis*

The sequence data were analyzed by the ProtParam (http://expasy.org/tools/protparam.html). Open-reading frame (ORF) of the DNA sequence was searched using the ORF finder software

(http://www.ncbi.nlm.nih.gov/gorf/gorf.html). The hydropathy analysis of the wheat TaZFP-1 amino acid sequence was generated by Protscale software (http://expasy.org/tools/protscale.html). Protein structure of the TaZFP-1 was deduced using the PredictProtein software (http://cubic.bioc.columbia.edu/ predictprotein/). Prediction of the TaZFP-1 sequence was simulated by the SWISS-MODEL software (http://swiss-model.expasy.org/). The domain of TaZFP-1 was prediction using SMART v4.0. Homology research of the wheat TaZFP-1 compared with the gene sequences of other species was performed using Blast 2.1 (http://www.ncbi.nlm.nih.gov/blast/), and a multiple sequence alignment was created with DNAMAN (version 5.2.2). Subsequently, the homology tree was also generated by DNAMAN (version 5.2.2).

## 3. Results

### *3.1 Analysis of the cDNA of TaZFP-1 from the wheat*

About 750 bp of the cDNA fragment was amplified using the pimers TaZFP-1 forward and TaZFP-1 reverse from the wheat (Figure 1). The sequence result indicated that the length of the cDNA was 759 bp. The TaZFP-1 cDNA sequence was submitted to GenBank (GenBank accession No. JF727656). The open reading frame happens to be 759 bp, coding a deduced protein of 252 amino acids (Figure 2). Primary structure analysis using ProtParam revealed that the molecular weight of the putative TaZFP-1 protein is 27.9 kD with a theoretical PI=6.23.

### *3.2 Hydrophobicity/hydrophilicity prediction of TaZFP-1 protein*

Hydrophobicity/hydrophilicity of TaZFP-1 protein was predicted utilizing Program of ProtScale. The results showed that the site of 159 (Ile) is the most hydrophobic (Score: 1.844) and 120 (Arg) the most hydrophilic (Score: -3.222) (Figure 3). Most amino acids in TaZFP-1 protein are hydrophilic amino acids, so the protein may be a soluble protein.

### *3.3 Secondary structure of TaZFP-1*

Predicted by PredictProtein Server program, the Secondary structure of TaZFP-1 is mainly composed of α-helices (H) (23.02%) and random coils (C) (59.13%), as well as extended strands (T) (1.98) and β-turns (E) (15.87%) (Figure 4).

### *3.4 Prediction and analysis functional domains of TaZFP-1*

The function domain of TaZFP-1 was predicted and analysis by SMART v4.0. The results (Figure 5a) showed that there is a RING finger domain. The RING finger domain, containing a C3HC4 (Cys3-His-Cys4) motif, is located at the C-terminus of TaZFP-1 from amino acid residues 196~233 (Figure 5b). One of the distinguished features of the TaZFP-1 RING domain is the presence of a zinc-binding motif containing two tetrahedrally coordinated zinc ions that stabilizes its globular structure (Figure 5c). Four pairs of metal-binding residues sequester two zinc ions at distinct tetrahedral sites forming a “cross-brace” motif (Figure 1d). In this motif, the first and third pairs bind one zinc ion, while the second and fourth pairs bind the other. The stability of the RING domain is a vital requirement for TaZFP-1 proteins to function as E3 ligases.

### *3.5 Homology analysis of TaZFP-1*

NCBI BLASTp showed that TaZFP-1 is homology to other plants RING finger protein to some extent, such as *Arabidopsis thaliana* L.(NP_568701), *Glycine max* (L.) Merr. (ACU24158), *Medicago truncatula* Gaertn. (ACJ83847), *Oryza sativa* L. (NP_001063596), *Populus trichocarpa* Torr. (XP_002321288), *Ricinus communis* L. (XP_002532746), *Sorghum bicolor* (L.) Moench (XP_002462621), *Vitis vinifera* L. (XP_002263351) and *Zea mays* L. (NP_001148999). The alignment result showed that although the amino acids from different plants have some differences, the C-terminus of those proteins containing a C3HC4 motif (Figure 6). The homology tree was consistent with the alignment result (Figure 7). From the Figure 7 we can see clearly that the RING domain E3 ligases from 10 plants classified into two groups. The first group contains the dicotyledon such as: *A. thaliana*, *P. trichocarpa, R. communis, G. max, M. truncatula, V. vinifera*. The second group contains monocotyledon such as: *O. sativa,* TaZFP-1, *S. bicolor, Z. mays*. TaZFP-1 showed a much greater similarity (73%) to *O. sativa* RING domain E3 ligases.

## 4. Discussion

The RING domain was originally described by Freemont and colleagues (1991). The basic sequence expression of the canonical RING is Cys-X2-Cys-X(9-39)-Cys-X(1-3)-His-X(2-3)-Cys-X2-Cys-X(4-48)-Cys-X2-Cys (where X is any amino acid). Three-dimensional structures of RING domains revealed that its conserved cysteine and histidine residues are buried within the domain’s core, where they help maintain the overall structure through binding two atoms of zinc. Additional semiconserved residues are implicated either in forming

the domain's hydrophobic core or in recruiting other proteins. Unlike zinc fingers, the zinc coordination sites in a RING finger are interleaved, yielding a rigid, globular platform for protein-protein interaction (Barlow et al, 1994; Borden et al, 1995). Up to now, only a few RING domain E3s had been characterized functionally in plants.

In this study, an EST sequence was isolated from SSH cDNA library of wheat three pistils lines, the EST shares approximately 66% identity to maize RING ubiquitin protein ligase, we tentatively designated the EST as TaZFP-1. Real-time RT-PCR analysis proved that TaZFP-1 expressed at a higher level in young spikes of three pistils lines as compared with the normal lines (Yang et al, 2010). It can be speculated that the over-expression of TaZFP-1 might have important function in morphological construction of the three pistils trait. Using RT-PCR method, we obtain the open reading frame (ORF) of the TaZFP-1. Functional domains analysis indicated that the TaZFP-1 is a RING finger domain protein and containing a C3HC4 (Cys3-His-Cys4) motif. The alignment result showed that although the amino acids from different plants have some differences, the C-terminus of those proteins containing a C3HC4 motif. Non-conservative sequences of RING finger domain protein might be responsibility for the specific recognition of target proteins. The molecular evolution threes shows that the TaZFP-1 was clustered into the monocotyledon group, and high genetic relationship with *O. sativa* RING domain E3 ligases. It means that TaZFP-1 has the similar function with the *O. sativa* RING domain E3 ligases.

At present, the RING domain E3 ligases in wheat have not been reported. In this study we acquired the full length cDNA sequence of RING domain E3 ligases TaZFP-1 in wheat by RT-PCR, but the function of TaZFP-1 need further studies.

**References**

Barlow, P. N., Luisi, B., Milner, A., Elliott, M., & Everett, R. (1994). Structure of the C3HC4 domain by 1H-nuclear magnetic resonance spectroscopy. A new structural class of zinc-finger. *Journal of Molecular Biology*, 237, 201–211.

Borden, K. L., & Freemont, P. S. (1996). The RING finger domain: a recent example of a sequence-structure family. *Current Opinion in Structural Biology*, 6, 395–401. doi:10.1016/S0959-440X(96)80060-1

Borden, K. L., Boddy, M. N., Lally, J., O'Reilly, N. J., & Martin, S. (1995). The solution structure of the RING finger domain from the acute promyelocytic leukaemia proto-oncoprotein PML. *The EMBO Journal,* 14, 1532–1541.

Freemont, P. S., Hanson, I. M., & Trowsdale, J. (1991). A novel cysteine-rich sequence motif. *Cell*, 64, 483–484. doi:10.1016/0092-8674(91)90229-R.

Hanson, I. M., Poustka, A., & Trowsdale, J. (1991). New genes in the class II region of the human major histocompatibility complex. *Genomics*, 10, 417–424. doi:10.1016/0888-7543(91)90327-B.

Lovering, R., Hanson, I. M., Borden, K. L., Martin, S., O'Reilly, N. J., Evan, G. I., Rahman, D., Pappin, D. J., Trowsdale, J., & Freemont, P. S. (1993). Identification and preliminary characterization of a protein motif related to the zinc finger. *Proceedings of the National Academy of Science U. S. A.*, 90, 2112–2116.

Manickavelu, A., Kambara, K., Mishina, K., & Koba, T. (2007). An efficient method for purifying high quality RNA from wheat pistils. *Colloids Surf B Biointerfaces*, 54, 254–258. doi:10.1016/j.colsurfb.2006.10.024.

Moon, J., Parry G., & Estelle M. (2004). The Ubiquitin-proteasome pathway and plant development. *The Plant Cell*, 16, 3181–3195. doi: 10.1105/tpc.104.161220.

Passmore, L. A., & Barford, D. (2004). Getting into position: The catalytic mechanisms of protein ubiquitylation. *Biochemical Journal*, *379*, 513-525. doi: 10.1042/BJ20040198.

Peng, Z. S. (2003). A new mutation in wheat producing three pistils in a floret. *J. Agronomy & Crop Science*, 189, 270-272. doi:10.1046/j.1439-037X.2003.00040.x.

VanDemark, A. P., & Hill, C. P. (2002). Structural basis of ubiquitylation. *Current Opinion in Structural Biology*, *12*, 822-830.

Weissman, A. M. (2001). Themes and variations on ubiquitylation. *Molecular Cell Biology*, *2*, 169-178. doi: 10.1038/ 35056563.

Yang, Z. J., Peng, Z. S., Yang, H., Yang, J., Wei, S. H., & Cai P. (2011). Suppression Subtractive Hybridization Identified Differentially Expressed Genes in Pistil Mutations in Wheat. *Plant Molecular Biology Reporter*, 29, 431-439. doi: 10.1007/s11105-010-0249-2.

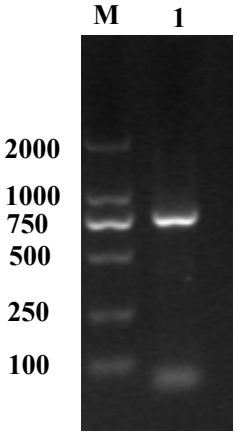


Figure 1. RT-PCR products of TaZFP-1

M: marker 2000; 1: RT-PCR PCR products

```
1   atgagtaataccagcggcgcacggcgtgccgcgaggaggcaatcgctggatggatctgctgataaagttgttgtggacttggatgcaagc
    M  S  N  T  S  G  A  R  R  A  A  R  R  Q  S  L  D  G  S  A  D  K  V  V  V  D  L  D  A  S
91  tcgccggtggcggggaaccatcatgggctctcggtcttccctggtcacgcacatcgcctattgatgtggaggctctcgatgatgaagtg
    S  P  V  A  G  N  H  H  G  L  S  V  F  P  G  A  R  T  S  P  I  D  V  E  A  L  D  D  E  V
181 caggcgctatcgccttcacaagtgcctcccccggggttgaacaggagaactagagggcaacctgtgacgccatttgctctggatgaagat
    Q  A  L  S  P  S  Q  V  P  P  P  G  L  N  R  R  T  R  G  Q  P  V  T  P  F  A  L  D  E  D
271 gctgttctggaagagaattggagaattataaggcaatgtgcgacaccatttttatctcgatgaagatgctgccatggaagagaatcggaga
    A  V  L  E  E  N  W  R  I  I  R  Q  C  A  T  P  F  Y  L  D  E  D  A  A  M  E  E  N  R  R
361 actagaagtcaatttttgacaccactttatctggatgaagatgctgccctggaagataatgcagctatgccttctcataacatgtggaac
    T  R  S  Q  F  L  T  P  L  Y  L  D  E  D  A  A  L  E  D  N  A  A  M  P  S  H  N  M  W  N
451 aaacgccaaagggttgcgcctttgatatgtctctccccggaaagggaagaagggtccagcctgcagtcgaacaatgcggtgcaaattagc
    K  R  Q  R  V  A  P  L  I  C  L  S  P  E  R  E  E  G  S  S  L  Q  S  N  N  A  V  Q  I  S
541 caagagcctgctaaggtggtggttccaaaggaaccgaatttcacttgcccggtgtgccttaacaagctggtggagccttcgacaacaaaa
    Q  E  P  A  K  V  V  V  P  K  E  P  N  F  T  C  P  V  C  L  N  K  L  V  E  P  S  T  T  K
631 tgtggccatatcttctgcgccgagtgcatcaagcaagccatccagtttcagaagaaatgccccacttgccgcaaggccctgaggaagaac
    C  G  H  I  F  C  A  E  C  I  K  Q  A  I  Q  F  Q  K  K  C  P  T  C  R  K  A  L  R  K  N
721 aacttccatcgtatttaccttccgaactcggatggttaa 759
    N  F  H  R  I  Y  L  P  N  S  D  G  *
```

Figure 2. The cDNA sequence and duduced amino acid of TaZFP-1


ISSN 1913-1844 E-ISSN 1913-1852

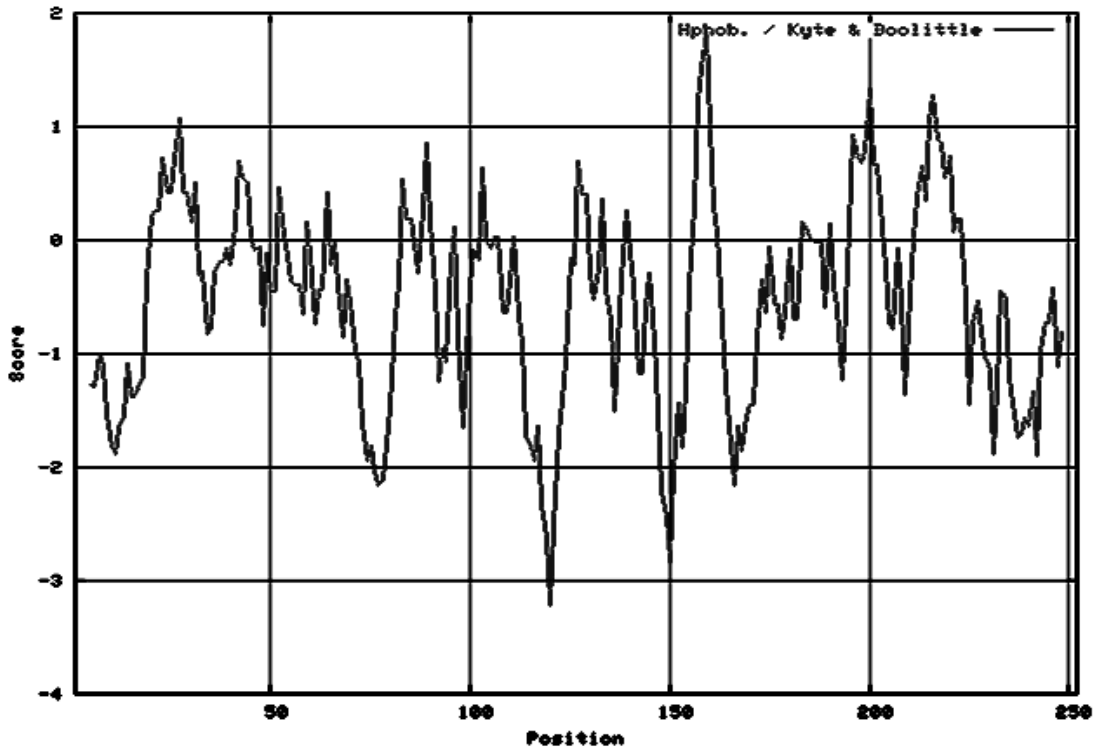


Figure 3. Hydrophilicity analysis of TaZFP-1


Figure 4. The secondary structure of TaZFP-1

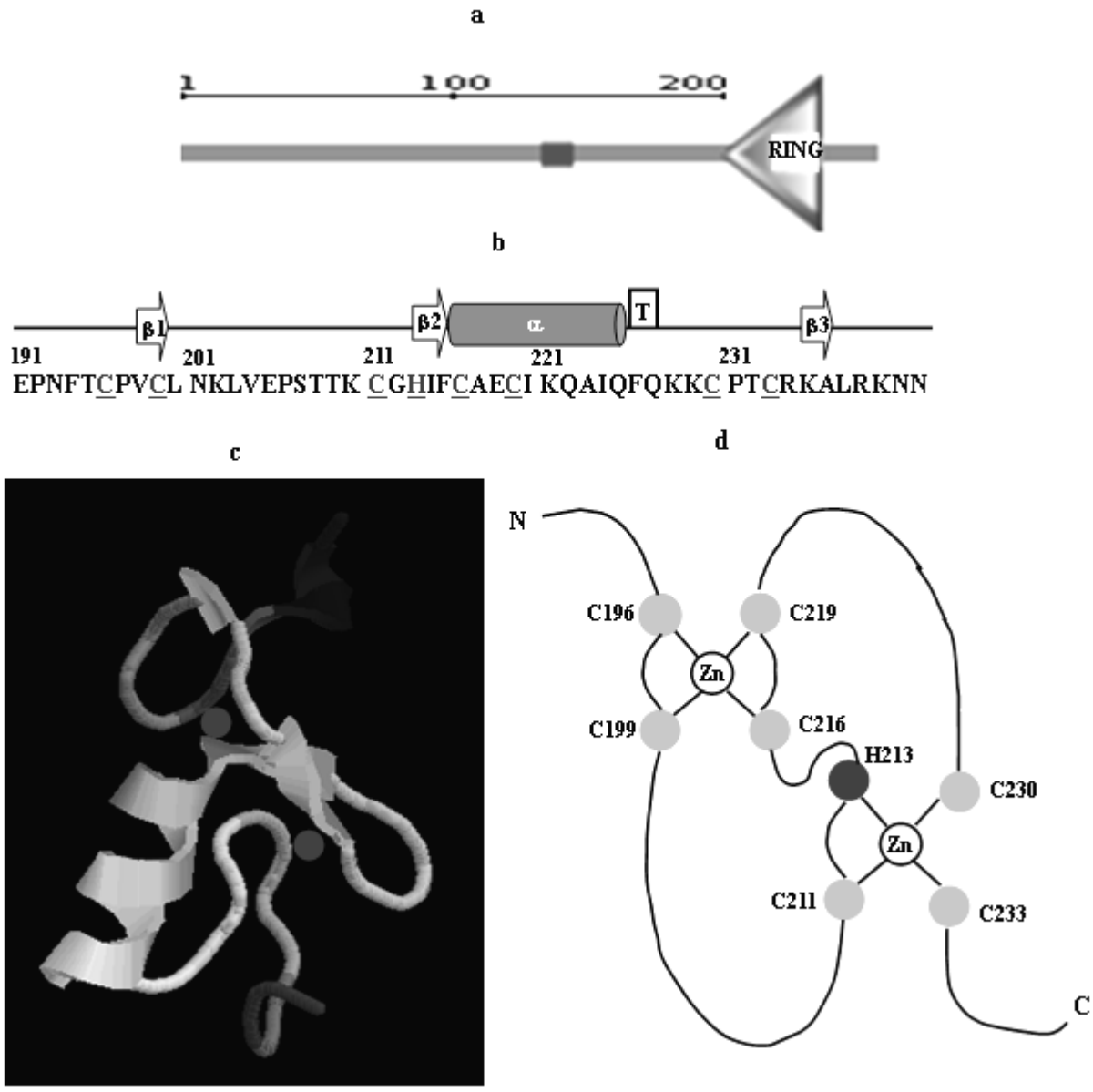


Figure 5. The domain analysis of TaZFP

a. The domain of TaZFP; b. Amino acid sequence and the secondary structural of TaZFP RING domain. The eight residues forming the C3HC4 are highlighted in underline; c. The 3D structue of TaZFP RING domain. The two zinc ions are show as spheres; d. Schematic representation of the ligation scheme employed by the TaZFP RING domain showing the unique "cross-brace" motif.

 ISSN 1913-1844 E-ISSN 1913-1852

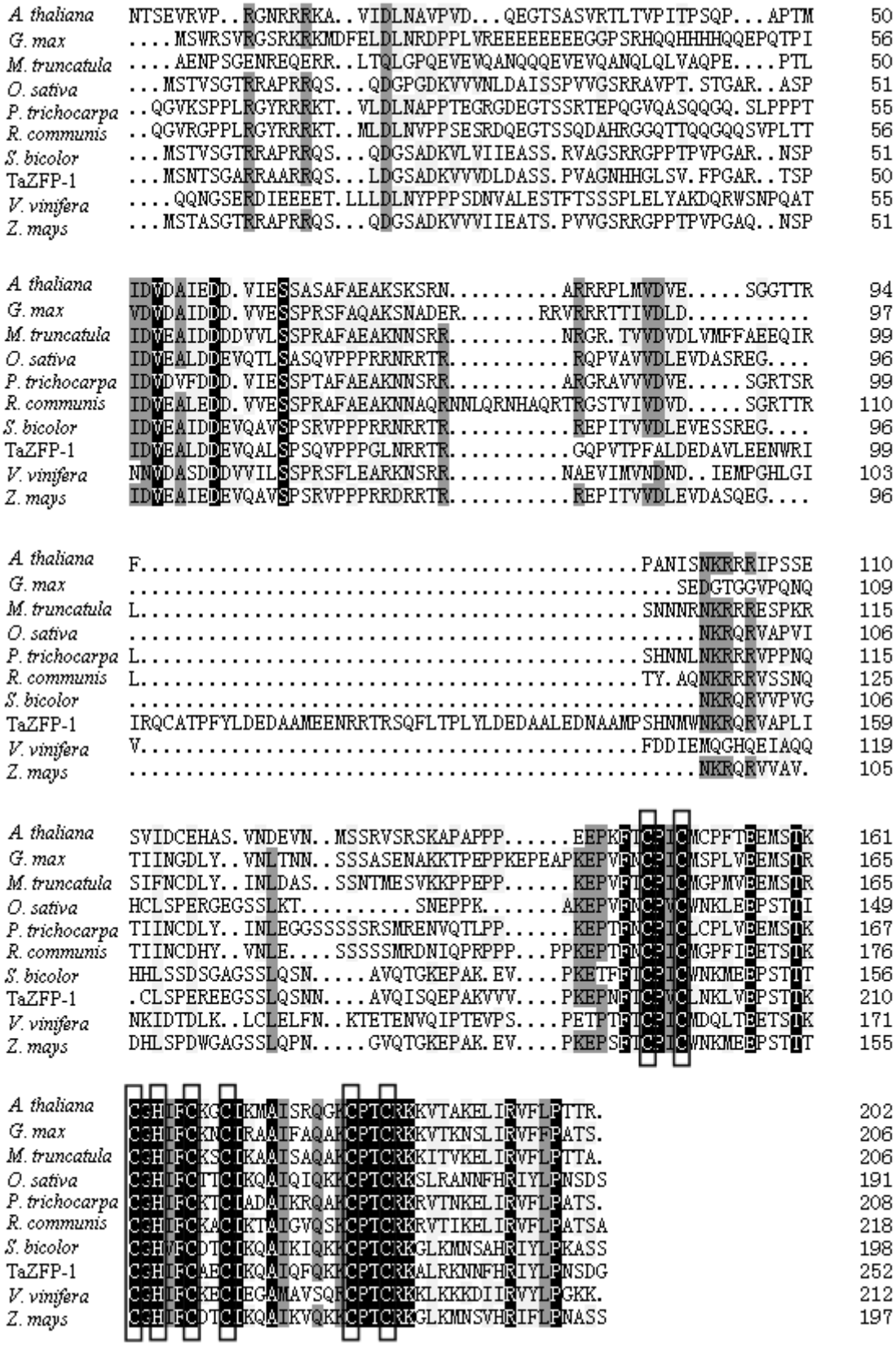


Figure 6. Comparison of the amino acid sequences of RING finger domain proteins among different plant
The eight residues forming the C3HC4 are marked with box.

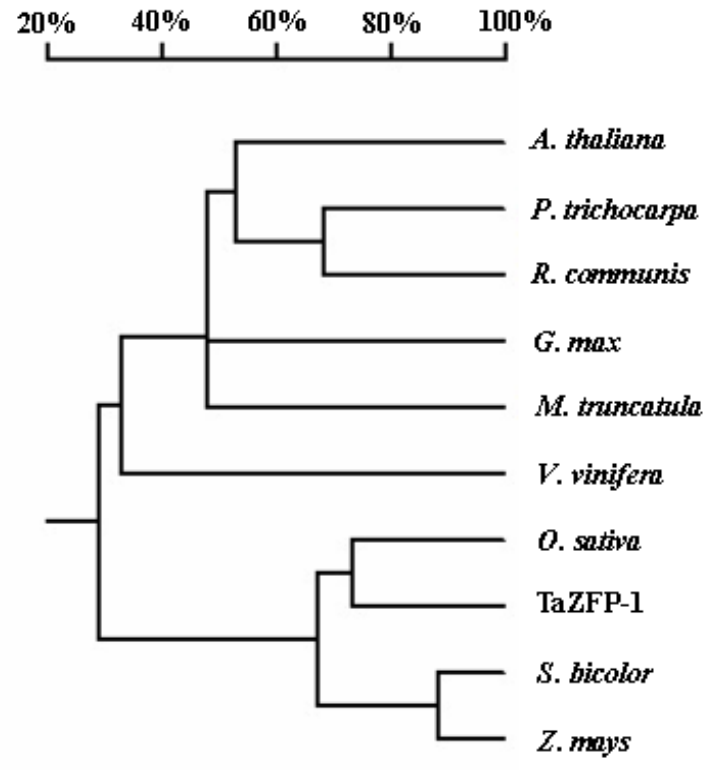


Figure 7. Polygenetic tree of RING finger domain proteins about 10 kinds of differents plants


ISSN 1913-1844 E-ISSN 1913-1852